18.175 Theory of Probability
Fall 2008

Section 7

Stopping times, Wald's identity. Another proof of SLLN.

Consider a sequence $(X_i)_{i\geq 1}$ of independent r.v.s and an integer valued random variable $V \in \{1, 2, \ldots\}$. We say that V is *independent of the future* if $\{V \leq n\}$ is independent of $\sigma((X_i)_{i\geq n+1})$. We say that V is a *stopping time* (Markov time) if $\{V \leq n\} \in \sigma(X_1, \ldots, X_n)$ for all n. Clearly, a stopping time is independent of the future. An example of stopping time is $V = \min\{k \geq 1, S_k \geq 1\}$.

Suppose that V is independent of the future. We can write

$$
\begin{aligned}
\mathbb{E}S_V &= \sum_{k\geq 1} \mathbb{E}S_V \mathrm{I}(V=k) = \sum_{k\geq 1} \mathbb{E}S_k \mathrm{I}(V=k) \\
&= \sum_{k\geq 1}\sum_{n\leq k} \mathbb{E}X_n \mathrm{I}(V=k) \overset{(*)}{=} \sum_{n\geq 1}\sum_{k\geq n} \mathbb{E}X_n \mathrm{I}(V=k) = \sum_{n\geq 1} \mathbb{E}X_n \mathrm{I}(V\geq n).
\end{aligned}
$$

In (*) we can interchange the order of summation if, for example, the double sequence is absolutely summable, by Fubini-Tonelli theorem. Since V is independent of the future, the event $\{V \geq n\} = \{V \leq n-1\}^c$ is independent of $\sigma(X_n)$ and we get

$$
\mathbb{E}S_V = \sum_{n\geq 1} \mathbb{E}X_n \times \mathbb{P}(V \geq n). \tag{7.0.1}
$$

This implies the following.

Theorem 14 *(Wald's identity.) If $(X_i)_{i\geq 1}$ are i.i.d., $\mathbb{E}|X_1| < \infty$ and $\mathbb{E}V < \infty$, then $\mathbb{E}S_V = \mathbb{E}X_1\mathbb{E}V$.*

Proof. By (7.0.1) we have,

$$
\mathbb{E}S_V = \sum_{n\geq 1} \mathbb{E}X_n\, \mathbb{P}(V\geq n) = \mathbb{E}X_1 \sum_{n\geq 1} \mathbb{P}(V \geq n) = \mathbb{E}X_1\, \mathbb{E}V.
$$

The reason we can interchange the order of summation in (*) is because under our assumptions the double sequence is absolutely summable since

$$
\sum_{n\geq 1}\sum_{k\geq n} \mathbb{E}|X_n|\mathrm{I}(V=k) = \sum_{n\geq 1} \mathbb{E}|X_n|\mathrm{I}(V\geq n) = \mathbb{E}|X_1|\,\mathbb{E}V < \infty,
$$

so we can apply Fubini-Tonelli theorem.

□

Theorem 15 *(Markov property) Suppose that $(X_i)_{i\geq 1}$ are i.i.d. and V is a stopping time. Then $(V, X_1, \ldots, X_V)$ is independent of $(X_{V+1}, X_{V+2}, \ldots)$ and*

$$
(X_{V+1}, X_{V+2}, \ldots) \overset{\text{dist}}{=} (X_1, X_2, \ldots),
$$

where $\stackrel{\text{dist}}{=}$ *means equality in distribution.*

Proof. Given a subset $N \subseteq \mathbb{N}$ and sequences (B_i) and (C_i) of Borel sets on $\mathbb{R}$, define events

$$A = \Big\{V \in N, X_1 \in B_1, ..., X_V \in B_V\Big\}$$

and for any $k \geq 1$,

$$D = \Big\{X_{V+1} \in C_1, ..., X_{V+k} \in C_k\Big\}.$$

We have,

$$\mathbb{P}(DA) = \sum_{n\geq 1} \mathbb{P}\left(DA\{V = n\}\right) = \sum_{n\geq 1} \mathbb{P}\left(D_n A\{V = n\}\right)$$

where

$$D_n = \{X_{n+1} \in C_1, \ldots, X_{n+k} \in C_k\}.$$

The intersection of events

$$A\{V = n\} = \begin{cases} \emptyset, & n \notin N \\ \{V = n, X_1 \in B_1, \ldots, X_n \in B_n\}, & \text{otherwise.} \end{cases}$$

Since V is a stopping time, $\{V = n\} \in \sigma(X_1, \ldots, X_n)$ and $A\{V = n\} \in \sigma(X_1, \ldots, X_n)$. On the other hand, $D_n \in \sigma(X_{n+1}, \ldots)$ and, as a result,

$$\mathbb{P}(DA) = \sum_{n\geq 1} \mathbb{P}(D_n)\mathbb{P}(A\{V = n\}) = \sum_{n\geq 1} \mathbb{P}(D_0)\mathbb{P}(A\{V = n\}) = \mathbb{P}(D_0)\mathbb{P}(A),$$

and this finishes the proof.

□

Remark. One could be a little bit more careful when talking about the events generated by a vector $(V, X_1, \ldots, X_V)$ that has random length. In the proof we implicitly assumed that such events are generated by events

$$A = \Big\{V \in N, X_1 \in B_1, ..., X_V \in B_V\Big\}$$

which is a rather intuitive definition. However, one could be more formal and define a σ-algebra of events generated by $(V, X_1, \ldots, X_V)$ as events A such that $A \cap \{V \leq n\} \in \sigma(X_1, \ldots, X_n)$ for any $n \geq 1$. This means that when $V \leq n$ the event A is expressed only in terms of $X_1, \ldots, X_n$. It is easy to check that with this more formal definition the proof remains exactly the same.

□

Let us give one interesting application of Markov property and Wald's identity that will yield another proof of strong law of large numbers.

Theorem 16 *Suppose that* $(X_i)_{i\geq 1}$ *are i.i.d. such that* $\mathbb{E}X_1 > 0$. *If* $Z = \inf_{n\geq 1} S_n$ *then* $\mathbb{P}(Z > -\infty) = 1$. *(Partial sums can not drift down to* $-\infty$ *if* $\mathbb{E}X_1 > 0$. *Of course, this is obvious by SLLN.)*

Proof. Let us define (see figure 7.1),

$$\tau_1 = \min\{k \geq 1, S_k \geq 1\},\ Z_1 = \min_{k\leq \tau_1} S_k,\ S_k^{(2)} = S_{\tau_1 + k} - S_{\tau_1},$$

$$\tau_2 = \min\Big\{k \geq 1, S_k^{(2)} \geq 1\Big\},\ Z_2 = \min_{k\leq \tau_2} S_k^{(2)},\ S_k^{(3)} = S_{\tau_2+k}^{(2)} - S_{\tau_2}^{(2)}.$$

By induction,

$$\tau_n = \min\Big\{k \geq 1, S_k^{(n)} \geq 1\Big\},\ Z_n = \min_{k\leq \tau_n} S_k^{(n)},\ S_k^{(n+1)} = S_{\tau_n+k}^{(n)} - S_{\tau_n}^{(n)}.$$

$Z_1, ..., Z_n$ are i.i.d. by Markov property.

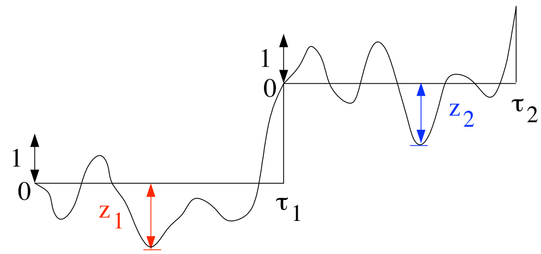

Figure 7.1: A sequence of stopping times.

Notice that, by construction, $S_{\tau_1+\ldots+\tau_{n-1}} \geq n-1$ and

$$Z = \inf_{k\geq 1} S_k = \inf\{Z_1, S_{\tau_1} + Z_2, S_{\tau_1+\tau_2} + Z_3, ...\}.$$

We have,

$$\{Z \leq -N\} = \bigcup_{k\geq 1}\{S_{\tau_1+\ldots+\tau_{k-1}} + Z_k \leq -N\} \subseteq \bigcup_{k\geq 1}\{k-1+Z_k \leq -N\}.$$

Therefore,

$$\begin{aligned}\mathbb{P}(Z \leq -N) &\leq \sum_{k\geq 1}\mathbb{P}(k-1+Z_k \leq -N) = \sum_{k\geq 1}\mathbb{P}(Z_k \leq -N-k+1)\\ &= \sum_{k\geq 1}\mathbb{P}(Z_1 \leq -N-k+1) = \sum_{j\geq N}\mathbb{P}(Z_1 \leq -j) \leq \sum_{j\geq N}\mathbb{P}(|Z_1| \geq j) \stackrel{N\to\infty}{\longrightarrow} 0\end{aligned}$$

if we can show that $\mathbb{E}|Z_1| < \infty$ since

$$\sum_{j\geq 1}\mathbb{P}(|Z_1| \geq j) \leq \mathbb{E}|Z_1| < \infty.$$

We can write

$$\mathbb{E}|Z_1| \leq \mathbb{E}\sum_{i\leq\tau_1}|X_i| \stackrel{\text{Wald}}{=} \mathbb{E}|X_1|\mathbb{E}\tau_1 < \infty$$

if we can show that $\mathbb{E}\tau_1 < \infty$. This is left as an exercise (hint: truncate X_i's and τ_1 and use Wald's identity). We proved that $\mathbb{P}(Z \leq -N) \stackrel{N\to\infty}{\longrightarrow} 0$ which, of course, implies that $\mathbb{P}(Z > -\infty) = 1$.

□

This result gives another proof of the SLLN.

Theorem 17 *If $(X_i)_{i\geq 1}$ are i.i.d. and $\mathbb{E}X_1 = 0$ then $\frac{S_n}{n} \to 0$ a.s.*

Proof. Given $\varepsilon > 0$ we define $X_i^\varepsilon = X_i + \varepsilon$ so that $\mathbb{E}X_1^\varepsilon = \varepsilon > 0$. By the above result, $\inf_{n\geq 1}(S_n + n\varepsilon) > -\infty$ with probability one. This means that for all $n \geq 1$, $S_n + n\varepsilon \geq -M > -\infty$ for some large enough M. Dividing both sides by n and letting $n \to \infty$ we get

$$\liminf_{n\to\infty}\frac{S_n}{n} \geq -\varepsilon$$

with probability one. We can then let $\varepsilon \to 0$ over some sequence. Similarly, we prove that $\limsup \frac{S_k}{k} \leq 0$ with probability one.

□